

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

- (i) APPLICANT: REGENTS OF THE UNIVERSITY OF MINNESOTA
- (ii) TITLE OF INVENTION: DNA MOLECULES AND PROTEIN DISPLAYING  
IMPROVED TRIAZINE COMPOUND DEGRADING ABILITY
- (iii) NUMBER OF SEQUENCES: 26
- (iv) CORRESPONDENCE ADDRESS:
  - (A) ADDRESSEE: MUETING, RAASCH & GEBHARDT, P.A.
  - (B) STREET: 119 North Fourth Street
  - (C) CITY: Minneapolis
  - (D) STATE: Minnesota
  - (E) COUNTRY: USA
  - (F) ZIP: 55401
- (v) COMPUTER READABLE FORM:
  - (A) MEDIUM TYPE: Floppy disk
  - (B) COMPUTER: IBM PC compatible
  - (C) OPERATING SYSTEM: PC-DOS/MS-DOS
  - (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
- (vi) PRIORITY APPLICATION DATA:
  - (A) APPLICATION NUMBER: 60/035,404
  - (B) FILING DATE: 17-JAN-1997
  - (C) CLASSIFICATION:
- (vii) CURRENT APPLICATION DATA:
  - (A) APPLICATION NUMBER: Not Assigned
  - (B) FILING DATE: 16-JAN-1998
  - (C) CLASSIFICATION:
- (viii) ATTORNEY/AGENT INFORMATION:
  - (A) NAME: MCCORMACK, MYRA M.
  - (B) REGISTRATION NUMBER: 36,602
  - (C) REFERENCE/DOCKET NUMBER: 110.00400201
- (ix) TELECOMMUNICATION INFORMATION:
  - (A) TELEPHONE: 612-305-1225
  - (B) TELEFAX: 612-305-1228

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

CTCGGGTAAC	TTCTTTGAGCG	CGGCCACAGC	AGCCTTGATC	ATGAAGGCGA	GCATGGTGAC	60
CTTGACGCCG	CTCTTTTCGT	TCTCTTTGTT	GAAC'TGCACG	CGAAAGGCTT	CCAGGTCGGT	120
GATGTCCGCG	TCGTCTGGT	TGGTGACGTG	CGGGATGACC	ACCCAGTTGC	GGTGCAGGTT	180
TTTCGATGGC	ATAATATCTG	CGTTGCGACG	TGTAACACAC	TATTGGAGAC	ATATCATGCA	240
AACGCTCAGC	ATCCAGCACG	GTACCCTCGT	CACGATGGAT	CAGTACCGCA	GAGTCCTTGG	300
GGATAGCTGG	GTTACAGTGC	AGGATGGACG	GATCGTCGCG	CTCGGAGTGC	ACGCCGAGTC	360
GGTGCCCTCCG	CCAGCGGATC	GGGTGATCGA	TGCACGCGGC	AAGGTCGTGT	TACCCGGTTT	420
CATCAATGCC	CACACCCATG	TGAACCAGAT	CCTCCTGCGC	GGAGGGCCCT	CGCACGGACG	480
TCAATTCTAT	GACTGGCTGT	TCAACGTTGT	GTATCCGGGA	CAAAAGGCGA	TGAGACCGGA	540
GGACGTAGCG	GTGGCGGTGA	GGTTGTATTG	TGCGGAAGCT	GTGCGCAGCG	GGATTACGAC	600
GATCAACGAA	AACGCCGATT	CGGCCATCTA	CCCAGGCAAC	ATCGAGGCCG	CGATGGCGGT	660
CTATGGTGAG	GTGGGTGTGA	GGGTCGTCTA	CGCCCGCATG	TTCTTTGATC	GGATGGACGG	720
GCGCATTCAA	GGGTATGTGG	ACGCCTTGAA	GGCTCGCTCT	CCCCAAGTCG	AACTGTGCTC	780
GATCATGGAG	GAAACGGCTG	TGGCCAAAGA	TCGGATCACA	GCCCTGTCAG	ATCAGTATCA	840
TGGCACGGCA	GGAGGTCGTA	TATCAGTTTG	GCCCGCTCCT	GCCACTACCA	CGGCGGTGAC	900
AGTTGAAGGA	ATGCGATGGG	CACAAGCCTT	CGCCCGTGAT	CGGGCGGTAA	TGTGGACGCT	960
TCACATGGCG	GAGAGCGATC	ATGATGAGCG	GATTCATGGG	ATGAGTCCCG	CCGAGTACAT	1020
GGAGTGTTAC	GGACTCTTGG	ATGAGCGTCT	GCAGGTCGCG	CATTGCGTGT	ACTTTGACCG	1080
GAAGGATGTT	CGGCTGCTGC	ACCGCCACAA	TGTGAAGGTC	GCGTCGCAGG	TTGTGAGCAA	1140
TGCCTACCTC	GGCTCAGGGG	TGGCCCCCGT	GCCAGAGATG	GTGGAGCGCG	GCATGGCCGT	1200
GGGCATTGGA	ACAGATAACG	GGAATAGTAA	TGACTCCGCA	AACATGATCG	GAGACATGAA	1260
GTTTATGGCC	CATATTACAC	GCGCGGTGCA	TCGGGATGCG	GACGTGCTGA	CCCCAGAGAA	1320

GATTCTTGAA ATGGCGACGA TCGATGGGGC GCGTTCGTTG GGAATGGACC ACGAGATTGG 1380  
 TTCCATCGAA ACCGGCAAGC GCGCGGACCT TATCCTGCTT GACCTGCGTC ACCTCAGACG 1440  
 ACTCTCACAT CATTTGGCGG CCACGATCGT GTTTCAGGCT TACGGCAATG AGGTGGACAC 1500  
 TGTCTGATT GACGGAAACG TTGTGATGGA GAACCGCCGC TTGAGCTTTC TTCCCCCTGA 1560  
 ACGTGAGTTG GCGTTCCTTG AGGAAGCGCA GAGCCGCGCC ACAGCTATTT TGCAGCGGGC 1620  
 GAACATGGTG GCTAACCAG CTTGGCGCAG CCTCTAGGAA ATGACGCCGT TGCTGCATCC 1680  
 GCCGCCCCCTT GAGGAAATCG CTGCCATCTT GGCGCGGCTC GGATTGGGGG GCGGACATGA 1740  
 CCTTGATGGA TACAGAATTG CCATGAATGC GGCACCTCCG TCCTTCGCTC GTGTGGAATC 1800  
 GTTGGTAGGT GAGGGTCGAC TCGGGGCGCC AGCTTCCCGA AGAGGTGAAA GGCCCCGAG 1858

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 473 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Gln	Thr	Leu	Ser	Ile	Gln	His	Gly	Thr	Leu	Val	Thr	Met	Asp	Gln
1				5					10					15	
Tyr	Arg	Arg	Val	Leu	Gly	Asp	Ser	Trp	Val	His	Val	Gln	Asp	Gly	Arg
			20					25					30		
Ile	Val	Ala	Leu	Gly	Val	His	Ala	Glu	Ser	Val	Pro	Pro	Pro	Ala	Asp
		35					40						45		
Arg	Val	Ile	Asp	Ala	Arg	Gly	Lys	Val	Val	Leu	Pro	Gly	Phe	Ile	Asn
		50					55				60				
Ala	His	Thr	His	Val	Asn	Gln	Ile	Leu	Leu	Arg	Gly	Gly	Pro	Ser	His
	65					70				75				80	
Gly	Arg	Gln	Phe	Tyr	Asp	Trp	Leu	Phe	Asn	Val	Val	Tyr	Pro	Gly	Gln
			85						90					95	
Lys	Ala	Met	Arg	Pro	Glu	Asp	Val	Ala	Val	Ala	Val	Arg	Leu	Tyr	Cys
			100						105					110	
Ala	Glu	Ala	Val	Arg	Ser	Gly	Ile	Thr	Thr	Ile	Asn	Glu	Asn	Ala	Asp
			115					120				125			

Ser	Ala	Ile	Tyr	Pro	Gly	Asn	Ile	Glu	Ala	Ala	Met	Ala	Val	Tyr	Gly
130						135					140				
Glu	Val	Gly	Val	Arg	Val	Val	Tyr	Ala	Arg	Met	Phe	Phe	Asp	Arg	Met
145					150					155					160
Asp	Gly	Arg	Ile	Gln	Gly	Tyr	Val	Asp	Ala	Leu	Lys	Ala	Arg	Ser	Pro
				165					170					175	
Gln	Val	Glu	Leu	Cys	Ser	Ile	Met	Glu	Glu	Thr	Ala	Val	Ala	Lys	Asp
			180					185					190		
Arg	Ile	Thr	Ala	Leu	Ser	Asp	Gln	Tyr	His	Gly	Thr	Ala	Gly	Gly	Arg
		195					200					205			
Ile	Ser	Val	Trp	Pro	Ala	Pro	Ala	Thr	Thr	Thr	Ala	Val	Thr	Val	Glu
	210					215					220				
Gly	Met	Arg	Trp	Ala	Gln	Ala	Phe	Ala	Arg	Asp	Arg	Ala	Val	Met	Trp
225					230					235					240
Thr	Leu	His	Met	Ala	Glu	Ser	Asp	His	Asp	Glu	Arg	Ile	His	Gly	Met
				245					250					255	
Ser	Pro	Ala	Glu	Tyr	Met	Glu	Cys	Tyr	Gly	Leu	Leu	Asp	Glu	Arg	Leu
			260					265					270		
Gln	Val	Ala	His	Cys	Val	Tyr	Phe	Asp	Arg	Lys	Asp	Val	Arg	Leu	Leu
		275					280					285			
His	Arg	His	Asn	Val	Lys	Val	Ala	Ser	Gln	Val	Val	Ser	Asn	Ala	Tyr
	290					295					300				
Leu	Gly	Ser	Gly	Val	Ala	Pro	Val	Pro	Glu	Met	Val	Glu	Arg	Gly	Met
305					310					315					320
Ala	Val	Gly	Ile	Gly	Thr	Asp	Asn	Gly	Asn	Ser	Asn	Asp	Ser	Ala	Asn
				325					330					335	
Met	Ile	Gly	Asp	Met	Lys	Phe	Met	Ala	His	Ile	His	Arg	Ala	Val	His
			340					345					350		
Arg	Asp	Ala	Asp	Val	Leu	Thr	Pro	Glu	Lys	Ile	Leu	Glu	Met	Ala	Thr
		355					360					365			
Ile	Asp	Gly	Ala	Arg	Ser	Leu	Gly	Met	Asp	His	Glu	Ile	Gly	Ser	Ile
	370					375					380				
Glu	Thr	Gly	Lys	Arg	Ala	Asp	Leu	Ile	Leu	Leu	Asp	Leu	Arg	His	Leu
385					390					395					400
Arg	Arg	Leu	Ser	His	His	Leu	Ala	Ala	Thr	Ile	Val	Phe	Gln	Ala	Tyr
				405					410					415	
Gly	Asn	Glu	Val	Asp	Thr	Val	Leu	Ile	Asp	Gly	Asn	Val	Val	Met	Glu
			420					425					430		

Asn Arg Arg Leu Ser Phe Leu Pro Pro Glu Arg Glu Leu Ala Phe Leu  
           435                                  440                                  445

Glu Glu Ala Gln Ser Arg Ala Thr Ala Ile Leu Gln Arg Ala Asn Met  
           450                                  455                                  460

Val Ala Asn Pro Ala Trp Arg Ser Leu  
           465                                  470

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1808 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GCGAGCATGG TGACCTTGAC GCCGCTCTTT TCGTTCCTCTT TGTGAACTG CACGCGAAAG	60
GCTTCCAGGT CGGTGATGTC CGCGTCGTCG TGGTTGGTGA CGTGCGGGAT GACCACCCAG	120
TTGCGGTGCA GGTTTTTCGA TGGCATAATA TCTGCGTTGC GACGTGTAAC AACTATTGG	180
AGACATATCA TGCAAACGCT CAGCATCCAG CACGGTACCC TCGTCACGAT GGATCAGTAC	240
CGCAGAGTCC TTGGGGATAG CTGGGTTCAC GTGCAGGATG GACGGATCGT CGCGCTCGGA	300
GTGCACGCCG AGTCGGTGCC TCCGCCAGCG GATCGGGTGA TCGATGCACG CGGCAAGGTC	360
GTGTTACCCG GTTTCATCAA TGCCACACC CATGTGAACC AGATCCTCCT GCGCGGAGGG	420
CCCTCGCACG GGCGTCAATT CTATGACTGG CTGTTCAACG TTGTGTATCC GGGACAAAAG	480
GCGATGAGAC CGGAGGACGT AGCGGTGGCG GTGAGGTTGT ATTGTGCGGA AGCTGTGCGC	540
AGCGGGATTA CGACGATCAA CGAAAACGCC GATTGCGCCA TCTACCCAGG CAACATCGAG	600
GCCGCGATGG CGGTCTATGG TGAGGTGGGT GTGAGGGTCG TCTACGCCCC CATGTTCTTT	660
GATCGGATGG ACGGGCGCAT TCAAGGGTAT GTGGACGCCT TGAAGGCTCG CTCTCCCCAA	720
GTCGAACTGT GCTCGATCAT GGAGGGAACG GCTGTGGCCA AAGATCGGAT CACAGCCCTG	780
TCAGATCAGT ATCATGGCAC GGCAGGAGGT CGTATATCAG TTTGGCCCGC TCCTGCCACT	840
ACCACGGCGG TGACAGTTGA AGGAATGCGA TGGGCACAAG CCTTCGCCCC TGATCGGGCG	900
GTAATGTGGA CGCTTCACAT GGCGGAGAGC GATCATGATG AGCGGATTCA TGGGATGAGT	960

CCCGCCGAGT ACATGGAGTG TTACGGACTC TTGGATGAGC GTCTGCAGGT CGCGCATTGC 1020  
 GTGTACTTTG ACCGGAAGGA TGTTTCGGCTG CTGCACCGCC ACAATGTGAA GGTTCGCGTCG 1080  
 CAGGTTGTGA GCAATGCCTA CCTCGGCTCA GGGGTGGCCC CCGTGCCAGA GATGGTGGAG 1140  
 CGCGGCATGG CCGTGGGCAT TGGAACAGAT AACGGGAATA GTAATGACTC CGTAAACATG 1200  
 ATCGGAGACA TGAAGTTTAT GGCCCATATT CACCGCGCGG TGCATCGGGA TCGGACGTG 1260  
 CTGACCCCAG AGAAGATTCT TGAAATGGCG ACGATCGATG GGGCGCGTTC GTTGGAATG 1320  
 GACCACGAGA TTGGTTCCAT CGAAACCGGC AAGCGCGCGG ACCTTATCCT GCTTGACCTG 1380  
 CGTCACCCTC AGACGACTCC TCACCATCAT TTGGCGGCCA CGATCGTGTT TCAGGCTTAC 1440  
 GGCAATGAGG TGGACACTGT CCTGATTGAC GGAAACGTTG TGATGGAGAA CCGCCGCTTG 1500  
 AGCTTTCTTC CCCCTGAACG TGAGTTGGCG TTCCTTGAGG AAGCGCAGAG CCGCGCCACA 1560  
 GCTATTTTGC AGCGGGCGAA CATGGTGGCT AACCAGCTT GGCGCAGCCT CTAGGAAATG 1620  
 ACGCCGTTGC TGCATCCGCC GCCCCTTGAG GAAATCGCTG CCATCTTGGC GCGGCTCGGA 1680  
 TTGGGGGGCG GACATGACCT TGATGGATAC AGAATTGCCA TGAATGCGGC ACTTCCGTCC 1740  
 TTCGCTCGTG TGAATCGTT GGTAGGTGAG GGTGACTGC GGGCGCCAGC TTCCGAAGA 1800  
 AGTGAAAG 1808

## (2) INFORMATION FOR SEQ ID NO:4:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1846 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GAGCGCCGCC ACAGCAGCCT TGATCATGAA GGCGAGCATG GTGACCTTGA CGCCGCTCTT 60  
 TTCGTTCTCT TTGTTGAACT GCACGCGAAA GGCTTCCAGG TCGGTGATGT CCGCGTCGTC 120  
 TGGGTTGGTG ACGTGCGGGA TGACCACCCA GTTGCGGTGC AGGTTTTTTCG ATGGCGTAAT 180  
 ATCTGCGTTG CGACGTGTAA CACACTATTG GAGACATATC ATGCAAACGC TCAGCATCCA 240  
 GCACGGTACC CTCGTCACGA TGGATCAGTA CCGCAGAGTC CTTGGGGATA GCTGGGTTCA 300  
 CGTGCAGGAT GGACGGATCG TCGCGCTCGG AGTGCACGCC GAGTCGGTGC CTCCGCCAGC 360

GGATCGGGTG ATCGATGCAC GCGGCAAGGT CGTGTTACCC GGTTCATCA ATGCCCACAC 420  
 CCATGTGAAC CAGATCCTCC TGC GCGGAGG GCCCTCGCAC GGGCGTCAAT TCTATGACTG 480  
 GCTGTTCAAC GTTGTGTATC CGGGACAAAA GGCGATGAGA CCGGAGGACG TAGCGGTGGC 540  
 GGTGAGGTTG TATTGTGCGG AAGCTGTGCG CAGCGGGATT ACGACGATCA ACGAAAACGC 600  
 CGATTGCGCC ATCTACCCAG GCAACATCGA GGCCGCGATG GCGGTCTATG GTGAGGTGGG 660  
 TGTGAGGGTC GTCTACGCCG GCATGTTCTT TGATCGGATG GACGGGCGCA TTCAAGGGTA 720  
 TGTGGACGCC TTGAAGGCTC GCTCTCCCCA AGTCGAACTG TGCTCGATCA TGGAGGAAAC 780  
 GGCTGTGGCC AAAGATCGGA TCACAGCCCT GTCAGATCAG TATCATGGCA CGGCAGGAGG 840  
 TCGTATATCA GTTTGGCCCG CTCCTGCCAC TACCACGGCG GTGACAGTTG AAGGAATGCG 900  
 ATGGGCACAA GCCTTCGCCC GTGATCGGGC GGTAATGTGG ACGCTTCACA TGGCGGAGAG 960  
 CGATCATGAT GAGCGGATTC ATGGGATGAG TCCCGCCGAT TACATGGAGT GTTACGGACT 1020  
 CTTGGATGAG CGTCTGCAGG TCGCGCATTG CGTGTA CTTT GACCGGAAGG ATGTTGCGCT 1080  
 GCTGCACCGC CACAATGTGA AGGTCGCGTC GCAGGTTGTG AGCAATGCCT ACCTCGGCTC 1140  
 AGGGGTGGCC CCCGTGCCAG AGATGGTGA GCGCGGCATG GCCGTGGGCA TTGGAACAGA 1200  
 TAACGGGAAT AGTAATGACT CCGTAAACAT GATCGGAGAC ATGAAGTTTA TGGCCCATAT 1260  
 TCACCGCGCG GTGCATCGGG ATGCGGACGT GCTGACCCCA GAGAAGATTC TTGAAATGGC 1320  
 GACGATCGAT GGGGCGCGTT CGTTGGGGAT GGACCACGAG ATTGGTTCCA TCGAAACCGG 1380  
 CAAGCGCGCG GACCTTATCC TGCTTGACCT GCGTCA CCGT CAGACGACTC CTCACCATCA 1440  
 TTTGGCGGCC ACGATCGTGT TTCAGGCTTA CGGCAATGAG GTGGACACTG TCCTGATTGA 1500  
 CGGAAACGTT GTGATGGAGA ACCGCCGCTT GAGCTTTCTT CCCCCTGAAC GTGAGTTGGC 1560  
 GTTCCTTGAG GAAGCGCAGA GCCGCGCCAC AGCTATTTTG CAGCGGGCGA ACATGGTGGC 1620  
 TAACCCAGCT TGGCGCAGCC TCTAGGAAAT GACGCCGTTG CTGCATCCGC CGCCCCTTGA 1680  
 GGAAATCGCT GCCATCTTGG CGCGGCTCGG ATTGGGGGGC GGACATGACC TTGATGGATA 1740  
 CAGAATTGCC ATGAATGCGG CACTTCCGTC CTTGCTCGT GTGGAATCGT TGGTAGGTGA 1800  
 GGGTCGACTG CGGGCGCCAG CTTCCCGAAG AAGTGAAAGG CCCGAG 1846

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 601 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Ala	Ser	Met	Val	Thr	Leu	Thr	Pro	Leu	Phe	Ser	Phe	Ser	Leu	Leu	Asn	1	5	10	15
Cys	Thr	Arg	Lys	Ala	Ser	Arg	Ser	Val	Met	Ser	Ala	Ser	Ser	Trp	Leu	20	25	30	
Val	Thr	Cys	Gly	Met	Thr	Thr	Gln	Leu	Arg	Cys	Arg	Phe	Phe	Asp	Gly	35	40	45	
Ile	Ile	Ser	Ala	Leu	Arg	Arg	Val	Thr	His	Tyr	Trp	Arg	His	Ile	Met	50	55	60	
Gln	Thr	Leu	Ser	Ile	Gln	His	Gly	Thr	Leu	Val	Thr	Met	Asp	Gln	Tyr	65	70	75	80
Arg	Arg	Val	Leu	Gly	Asp	Ser	Trp	Val	His	Val	Gln	Asp	Gly	Arg	Ile	85	90	95	
Val	Ala	Leu	Gly	Val	His	Ala	Glu	Ser	Val	Pro	Pro	Pro	Ala	Asp	Arg	100	105	110	
Val	Ile	Asp	Ala	Arg	Gly	Lys	Val	Val	Leu	Pro	Gly	Phe	Ile	Asn	Ala	115	120	125	
His	Thr	His	Val	Asn	Gln	Ile	Leu	Leu	Arg	Gly	Gly	Pro	Ser	His	Gly	130	135	140	
Arg	Gln	Phe	Tyr	Asp	Trp	Leu	Phe	Asn	Val	Val	Tyr	Pro	Gly	Gln	Lys	145	150	155	160
Ala	Met	Arg	Pro	Glu	Asp	Val	Ala	Val	Ala	Val	Arg	Leu	Tyr	Cys	Ala	165	170	175	
Glu	Ala	Val	Arg	Ser	Gly	Ile	Thr	Thr	Ile	Asn	Glu	Asn	Ala	Asp	Ser	180	185	190	
Ala	Ile	Tyr	Pro	Gly	Asn	Ile	Glu	Ala	Ala	Met	Ala	Val	Tyr	Gly	Glu	195	200	205	
Val	Gly	Val	Arg	Val	Val	Tyr	Ala	Arg	Met	Phe	Phe	Asp	Arg	Met	Asp	210	215	220	
Gly	Arg	Ile	Gln	Gly	Tyr	Val	Asp	Ala	Leu	Lys	Ala	Arg	Ser	Pro	Gln	225	230	235	240
Val	Glu	Leu	Cys	Ser	Ile	Met	Glu	Gly	Thr	Ala	Val	Ala	Lys	Asp	Arg	245	250	255	

005597 055500



Ile	Thr	Ala	Leu	Ser	Asp	Gln	Tyr	His	Gly	Thr	Ala	Gly	Gly	Arg	Ile
			260					265					270		
Ser	Val	Trp	Pro	Ala	Pro	Ala	Thr	Thr	Thr	Ala	Val	Thr	Val	Glu	Gly
		275					280					285			
Met	Arg	Trp	Ala	Gln	Ala	Phe	Ala	Arg	Asp	Arg	Ala	Val	Met	Trp	Thr
		290				295					300				
Leu	His	Met	Ala	Glu	Ser	Asp	His	Asp	Glu	Arg	Ile	His	Gly	Met	Ser
305					310					315					320
Pro	Ala	Glu	Tyr	Met	Glu	Cys	Tyr	Gly	Leu	Leu	Asp	Glu	Arg	Leu	Gln
				325				330						335	
Val	Ala	His	Cys	Val	Tyr	Phe	Asp	Arg	Lys	Asp	Val	Arg	Leu	Leu	His
			340					345					350		
Arg	His	Asn	Val	Lys	Val	Ala	Ser	Gln	Val	Val	Ser	Asn	Ala	Tyr	Leu
		355					360					365			
Gly	Ser	Gly	Val	Ala	Pro	Val	Pro	Glu	Met	Val	Glu	Arg	Gly	Met	Ala
		370				375					380				
Val	Gly	Ile	Gly	Thr	Asp	Asn	Gly	Asn	Ser	Asn	Asp	Ser	Val	Asn	Met
385					390					395				400	
Ile	Gly	Asp	Met	Lys	Phe	Met	Ala	His	Ile	His	Arg	Ala	Val	His	Arg
				405					410					415	
Asp	Ala	Asp	Val	Leu	Thr	Pro	Glu	Lys	Ile	Leu	Glu	Met	Ala	Thr	Ile
			420					425					430		
Asp	Gly	Ala	Arg	Ser	Leu	Gly	Met	Asp	His	Glu	Ile	Gly	Ser	Ile	Glu
		435				440						445			
Thr	Gly	Lys	Arg	Ala	Asp	Leu	Ile	Leu	Leu	Asp	Leu	Arg	His	Pro	Gln
		450				455					460				
Thr	Thr	Pro	His	His	His	Leu	Ala	Ala	Thr	Ile	Val	Phe	Gln	Ala	Tyr
465					470					475					480
Gly	Asn	Glu	Val	Asp	Thr	Val	Leu	Ile	Asp	Gly	Asn	Val	Val	Met	Glu
				485					490					495	
Asn	Arg	Arg	Leu	Ser	Phe	Leu	Pro	Pro	Glu	Arg	Glu	Leu	Ala	Phe	Leu
			500					505					510		
Glu	Glu	Ala	Gln	Ser	Arg	Ala	Thr	Ala	Ile	Leu	Gln	Arg	Ala	Asn	Met
		515					520					525			
Val	Ala	Asn	Pro	Ala	Trp	Arg	Ser	Leu	Glu	Met	Thr	Pro	Leu	Leu	His
		530				535					540				
Pro	Pro	Pro	Leu	Glu	Glu	Ile	Ala	Ala	Ile	Leu	Ala	Arg	Leu	Gly	Leu
545					550					555					560

(2) INFORMATION FOR SEQ ID NO:6:

(A) LENGTH: 614 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Ser 1	Ala	Ala	Thr 5	Ala	Leu	Ile	Met 10	Lys	Ala	Ser	Met	Val 15	Thr	Leu
Thr	Pro	Leu	Phe 20	Ser	Phe	Ser	Leu	Leu	Asn	Cys	Thr	Arg 30	Lys	Ala
Arg	Ser	Val 35	Met	Ser	Ala	Ser	Ser	Trp	Leu	Val	Thr	Cys 45	Gly	Met
Thr 50	Gln	Leu	Arg	Cys	Arg	Phe 55	Phe	Asp	Gly	Val	Ile 60	Ser	Ala	Leu
Arg 65	Val	Thr	His	Tyr	Trp 70	Arg	His	Ile	Met	Gln 75	Thr	Leu	Ser	Ile
His	Gly	Thr	Leu 85	Val	Thr	Met	Asp	Gln	Tyr 90	Arg	Arg	Val	Leu	Gly
Ser	Trp	Val	His 100	Val	Gln	Asp	Gly	Arg	Ile 105	Val	Ala	Leu	Gly	Val
Ala	Glu	Ser	Val	Pro	Pro	Pro	Ala	Asp	Arg	Val	Ile	Asp 125	Ala	Arg
Lys 130	Val	Val	Leu	Pro	Gly	Phe 135	Ile	Asn	Ala	His	Thr 140	His	Val	Asn
Ile 145	Leu	Leu	Arg	Gly	Gly 150	Pro	Ser	His	Gly	Arg 155	Gln	Phe	Tyr	Asp
Leu	Phe	Asn	Val	Val	Tyr	Pro	Gly	Gln	Lys 170	Ala	Met	Arg	Pro	Glu

Val	Ala	Val	Ala	Val	Arg	Leu	Tyr	Cys	Ala	Glu	Ala	Val	Arg	Ser	Gly	
			180					185					190			
Ile	Thr	Thr	Ile	Asn	Glu	Asn	Ala	Asp	Ser	Ala	Ile	Tyr	Pro	Gly	Asn	
		195					200					205				
Ile	Glu	Ala	Ala	Met	Ala	Val	Tyr	Gly	Glu	Val	Gly	Val	Arg	Val	Val	
	210					215					220					
Tyr	Ala	Arg	Met	Phe	Phe	Asp	Arg	Met	Asp	Gly	Arg	Ile	Gln	Gly	Tyr	
225					230					235					240	
Val	Asp	Ala	Leu	Lys	Ala	Arg	Ser	Pro	Gln	Val	Glu	Leu	Cys	Ser	Ile	
				245					250					255		
Met	Glu	Glu	Thr	Ala	Val	Ala	Lys	Asp	Arg	Ile	Thr	Ala	Leu	Ser	Asp	
			260					265					270			
Gln	Tyr	His	Gly	Thr	Ala	Gly	Gly	Arg	Ile	Ser	Val	Trp	Pro	Ala	Pro	
		275					280					285				
Ala	Thr	Thr	Thr	Ala	Val	Thr	Val	Glu	Gly	Met	Arg	Trp	Ala	Gln	Ala	
	290					295					300					
Phe	Ala	Arg	Asp	Arg	Ala	Val	Met	Trp	Thr	Leu	His	Met	Ala	Glu	Ser	
305					310					315					320	
Asp	His	Asp	Glu	Arg	Ile	His	Gly	Met	Ser	Pro	Ala	Asp	Tyr	Met	Glu	
				325					330					335		
Cys	Tyr	Gly	Leu	Leu	Asp	Glu	Arg	Leu	Gln	Val	Ala	His	Cys	Val	Tyr	
			340					345					350			
Phe	Asp	Arg	Lys	Asp	Val	Arg	Leu	Leu	His	Arg	His	Asn	Val	Lys	Val	
		355					360					365				
Ala	Ser	Gln	Val	Val	Ser	Asn	Ala	Tyr	Leu	Gly	Ser	Gly	Val	Ala	Pro	
	370					375					380					
Val	Pro	Glu	Met	Val	Glu	Arg	Gly	Met	Ala	Val	Gly	Ile	Gly	Thr	Asp	
385					390					395					400	
Asn	Gly	Asn	Ser	Asn	Asp	Ser	Val	Asn	Met	Ile	Gly	Asp	Met	Lys	Phe	
				405					410					415		
Met	Ala	His	Ile	His	Arg	Ala	Val	His	Arg	Asp	Ala	Asp	Val	Leu	Thr	
			420					425					430			
Pro	Glu	Lys	Ile	Leu	Glu	Met	Ala	Thr	Ile	Asp	Gly	Ala	Arg	Ser	Leu	
		435					440					445				
Gly	Met	Asp	His	Glu	Ile	Gly	Ser	Ile	Glu	Thr	Gly	Lys	Arg	Ala	Asp	
	450					455					460					

Leu Ile Leu Leu Asp Leu Arg His Pro Gln Thr Thr Pro His His His  
 465 470 475 480  
 Leu Ala Ala Thr Ile Val Phe Gln Ala Tyr Gly Asn Glu Val Asp Thr  
 485 490 495  
 Val Leu Ile Asp Gly Asn Val Val Met Glu Asn Arg Arg Leu Ser Phe  
 500 505 510  
 Leu Pro Pro Glu Arg Glu Leu Ala Phe Leu Glu Glu Ala Gln Ser Arg  
 515 520 525  
 Ala Thr Ala Ile Leu Gln Arg Ala Asn Met Val Ala Asn Pro Ala Trp  
 530 535 540  
 Arg Ser Leu Glu Met Thr Pro Leu Leu His Pro Pro Pro Leu Glu Glu  
 545 550 555 560  
 Ile Ala Ala Ile Leu Ala Arg Leu Gly Leu Gly Gly Gly His Asp Leu  
 565 570 575  
 Asp Gly Tyr Arg Ile Ala Met Asn Ala Ala Leu Pro Ser Phe Ala Arg  
 580 585 590  
 Val Glu Ser Leu Val Gly Glu Gly Arg Leu Arg Ala Pro Ala Ser Arg  
 595 600 605  
 Arg Ser Glu Arg Pro Glu  
 610

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 545 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGGTATCGGG GAATTCCTGA GCGCGGCCAC AGCAGCCNTG ATCATGAAGG CGAGCATGGT	60
GACCTNGACG CCGTNTTTTN GTTNTTTTTT GTTGAAGTGC ACGCGAAAGG TTCCAGGTGC	120
GTGATGTCCG CGTCGTCGTG GTTGGTGACG TGCGGGATGA CCACCCAGNT GCGGTGCAGG	180
TTTTTCGATG GCATAATATC TGC GTTGCGA CGTGTAACAC ACTANTGGAG ACATATCATG	240
CAAACGCTCA GCATCCAGCA CGGTACCCTC GTCACGATGG ATCAGTACCG CAGAGTCCTT	300
GGGGATAGCT GGGTTCACGT GCAGGATGGA CGGATCGTCG CGCTCGGAGT GCACGCCGAG	360

TCGGTGCCTC CGCCAGCGGA TCGGGTGATC GATGCACGCG GCAAGGTCGT GTTACCCGGT 420  
 TTCATCAATG CCCACACCCA TGTGAACCAG ATCCTCCTGC GCGGAGGGCC CTCGCACGGG 480  
 CGTCAATTCT ATGACTGGCT GTTCAACGTT GTGTATCCGG GACAAAAGGC GATGAGACCG 540  
 GAGGA 545

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 499 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CCTGCGCGGA GGGCCTCCGC ACGGGCGTCA ATTCTATGAC TGGCTGTTCA ACGTTGTGTA 60  
 TCCGGGACAA AAGGCGATGA GACCGGAGGA CGTAGCGGTG GCGGTGAGGT TGTATTGTGC 120  
 GGAAGCTGTG CGCAGCGGGA TTACGACGAT CAACGAAAAC GCCGATTCGG CCATCTACCC 180  
 AGGCAACATC GAGGCCGCGA TGGCGGTCTA TGGTGAGGTG GGTGTGAGGG TCGTCTACGC 240  
 CCGCATGTTT TTTGATCGGA TGGACGGGCG CATTCAAGGG TATGTGGACG CCTTGAAGGC 300  
 TCGCTCTCCC CAAGTCGAAC TGTGCTCGAT CATGGAGGAA ACGGCTGTGG CCAAAGATCG 360  
 GATCACAGCC CTGTCAGATC AGTATCATGG CACGGCAGGA GGTCCTATAT CAGTTTGGCC 420  
 CGCTCCTGCC ACTACCACGG CGGTGACATT TAAANGAATC CATGGGCCAA CCTCCCCCGT 480  
 GATCCGGCGG TAATGTGAC 499

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 360 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TNGCAGGTTG TGAGCATGCT ACTTCGGTTC AGNGTGGCC CCCGTGCCAG AGATGGTGGA 60  
 GCGCGGCATG GCCGTGGGCA TTGGAACAGA TAACGGGAAT AGTAATGACT CCGTAAACAT 120  
 GATCGGAGAC ATGAAGTTTA TGGCCCATAT TCACCGCGCG GTGCATCGGG ATGCGGACGT 180  
 GCTGACCCCA GAGAAGATTN TTGAAATGGC GACGATCGAT GGGGCGCGTT TCGTTGGGGA 240  
 TGGACCACGA GATTGGTTCC ATCGAAACCG GCAAGCGCGC GGACCTTATC CTGCTTGACC 300  
 TCGGTCACCC TCAGACGACT CCTCACCATC ATTTGGCGGC CACGATCGTG TTTCAGGCTT 360

## (2) INFORMATION FOR SEQ ID NO:10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 443 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CGGCCACGAT CGTGTTTCAG GCTTACGGCA ATGAGGTGGA CACTGTCCTG ATTGACGGAA 60  
 ACGTTGTGAT GGAGAACCGC CGCTTGAGCT TTCTTCCCCC TGAACGTGAG TTGGCGTTCC 120  
 TTGAGGAAGC GCAGAGCCGC GCCACAGCTA TTTTGCATCG GCGGAAACAT GGTGGCTAAC 180  
 CCAGCTTGGC GCAGCCTCTA GGAAATGACG CCGTTGCTGC ATCCGCCGCC CCTTGAGGAA 240  
 ATCGCTGCCA TCTTGGCGCG GCTCGGATTG GGGGGCGGAC ATGACCTTGA TGGATACAGA 300  
 ATTGCCATGA ATGCGGCACT TCCGTCCTTC GCTCGTGTGG AATCGTTGGT AGGTGAGGGT 360  
 CGACTGCGGG CGCCAGCTTC CCGAAGAGGT GAAAGCCCGA GGATCCTCTA GAGTCCGATT 420  
 TTTCCGATGT CATCACCGGC GCG 443

## (2) INFORMATION FOR SEQ ID NO:11:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 505 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CCTGCGCGGA GGCCTCCGCA CGGGCGTCAA TTCTATGACT GGCTGTTCAA CGTTGTGTAT 60  
 CCGGGACAAA AGGCGATGAG ACCGGAGGAC GTANCGGTGG CCGTGAGGTT GTATTGTGCG 120  
 GAAGCTGTGC GCAGCGGGAT TACGACGATC AACGAAAACG CCGATTCTGGC CATCTACCCA 180  
 GGCAACATCG AGGCCGCGAT GCGGCTCTAT GGTGAGGTGG GTGTGAGGGT CGTCTACGCC 240  
 CGCATGTTCT TTGATCGGAT GGACGGGCGC ATTCAAGGGT ATGTGGACGC CTTGAAGGCT 300  
 CGCTCTCCCC AAGTCGAACT GTGCTCGATC ATGGAGGAAA CGGCTGTGGC CAAAGATCGG 360  
 ATCACANCCC TGTCAGATCA NTATCATGGC ACGGCANGAG GTCCTATATC ANTTTGGCCC 420  
 GCTCCTGCCA CTACCACNGC GGTGACATTT NAANGAATTC CATNGGCACA ACCTTCCCCC 480  
 GTGATCNGGC GGTAATGTNG ACCCA 505

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 144 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Pro His Gly Arg Gln Phe Tyr Asp Trp Leu Phe Asn Val Leu Tyr Pro  
 1 5 10 15  
 Gly Gln Lys Ala Met Arg Pro Glu Asp Val Ala Val Ala Val Arg Leu  
 20 25 30  
 Tyr Cys Ala Glu Ala Val Arg Ser Gly Ile Thr Thr Ile Asn Glu Asn  
 35 40 45  
 Ala Asp Ser Ala Ile Tyr Pro Gly Asn Ile Glu Ala Ala Met Ala Val  
 50 55 60  
 Tyr Gly Glu Val Gly Val Arg Val Val Tyr Ala Arg Met Phe Phe Asp  
 65 70 75 80  
 Arg Met Asp Gly Arg Ile Gln Gly Tyr Val Asp Ala Leu Lys Ala Arg  
 85 90 95  
 Ser Pro Gln Val Glu Leu Cys Ser Ile Met Glu Glu Thr Ala Val Ala  
 100 105 110

Lys Asp Arg Ile Thr Ala Leu Ser Asp Gln Tyr His Gly Thr Ala Gly  
 115 120 125

Gly Arg Ile Ser Val Trp Pro Ala Pro Ala Thr Thr Thr Ala Val Thr  
 130 135 140

## (2) INFORMATION FOR SEQ ID NO:13:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 144 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Ser His Gly Arg Gln Phe Tyr Asp Trp Leu Phe Asn Val Leu Tyr Pro  
 1 5 10 15

Gly Gln Lys Ala Met Arg Pro Glu Asp Val Ala Val Ala Val Arg Leu  
 20 25 30

Tyr Cys Ala Glu Ala Val Arg Ser Gly Ile Thr Thr Ile Asn Glu Asn  
 35 40 45

Ala Asp Ser Ala Ile Tyr Pro Gly Asn Ile Glu Ala Ala Met Ala Val  
 50 55 60

Tyr Gly Glu Val Gly Val Arg Val Val Tyr Ala Arg Met Phe Phe Asp  
 65 70 75 80

Arg Met Asp Gly Arg Ile Gln Gly Tyr Val Asp Thr Leu Lys Ala Arg  
 85 90 95

Ser Pro Gln Val Glu Leu Cys Ser Ile Met Glu Glu Thr Ala Val Ala  
 100 105 110

Lys Asp Arg Ile Thr Ala Leu Ser Asp Gln Tyr His Gly Thr Ala Gly  
 115 120 125

Gly Arg Ile Ser Val Trp Pro Ala Pro Ala Thr Thr Thr Ala Val Thr  
 130 135 140

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 144 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single



(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Pro	His	Gly	Arg	Gln	Phe	Tyr	Asp	Trp	Leu	Phe	Asn	Val	Val	Tyr	Pro	1	5	10	15
Gly	Gln	Lys	Ala	Met	Arg	Pro	Glu	Asp	Val	Ala	Val	Ala	Val	Arg	Leu	20	25	30	
Tyr	Cys	Ala	Glu	Ala	Val	Arg	Ser	Gly	Ile	Thr	Thr	Ile	Asn	Glu	Asn	35	40	45	
Ala	Asp	Ser	Ala	Ile	Tyr	Pro	Gly	Asn	Ile	Glu	Ala	Ala	Met	Ala	Val	50	55	60	
Tyr	Gly	Glu	Val	Gly	Val	Arg	Val	Val	Tyr	Ala	Arg	Met	Phe	Phe	Asp	65	70	75	80
Arg	Met	Asp	Gly	Arg	Ile	Gln	Gly	Tyr	Val	Asp	Ala	Leu	Lys	Ala	Arg	85	90	95	
Ser	Pro	Gln	Val	Glu	Leu	Cys	Ser	Ile	Met	Glu	Glu	Thr	Ala	Val	Ala	100	105	110	
Lys	Asp	Arg	Ile	Thr	Ala	Leu	Ser	Asp	Gln	Tyr	His	Gly	Thr	Ala	Gly	115	120	125	
Gly	Arg	Ile	Ser	Val	Trp	Pro	Ala	Pro	Ala	Thr	Thr	Thr	Ala	Val	Thr	130	135	140	

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 145 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Ser	His	Gly	Arg	Gln	Phe	Tyr	Asp	Trp	Leu	Phe	Asn	Val	Leu	Tyr	Pro	1	5	10	15
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	---	---	----	----

Gly Gln Lys Ala Met Arg Pro Glu Asp Val Ala Val Ala Val Arg Leu  
 20 25 30  
 Tyr Cys Ala Glu Ala Val Arg Ser Gly Ile Thr Thr Ile Asn Glu Asn  
 35 40 45  
 Asn Ala Asp Ser Ala Ile Tyr Pro Gly Asn Ile Glu Ala Ala Met Ala  
 50 55 60  
 Val Tyr Gly Glu Val Gly Val Arg Val Val Tyr Ala Arg Met Phe Phe  
 65 70 75 80  
 Asp Arg Met Asp Gly Arg Ile Gln Gly Tyr Val Asp Thr Leu Lys Ala  
 85 90 95  
 Arg Ser Pro Gln Val Glu Leu Cys Ser Ile Met Glu Glu Thr Ala Val  
 100 105 110  
 Ala Lys Asp Arg Ile Thr Ala Leu Ser Asp Gln Tyr His Gly Thr Ala  
 115 120 125  
 Gly Gly Arg Ile Ser Val Trp Pro Ala Pro Ala Thr Thr Thr Ala Val  
 130 135 140  
 Thr  
 145

## (2) INFORMATION FOR SEQ ID NO:16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 144 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Ser His Gly Arg Gln Phe Tyr Asp Trp Leu Phe Asn Val Val Tyr Pro  
 1 5 10 15  
 Gly Gln Lys Ala Met Arg Pro Glu Asp Val Ala Val Ala Val Arg Leu  
 20 25 30  
 Tyr Cys Ala Glu Ala Val Arg Ser Gly Ile Thr Thr Ile Asn Glu Asn  
 35 40 45  
 Ala Asp Ser Ala Ile Tyr Pro Gly Asn Ile Glu Ala Ala Met Ala Val  
 50 55 60  
 Tyr Gly Glu Val Gly Val Arg Val Val Tyr Ala Arg Met Phe Phe Asp  
 65 70 75 80

CGCGAAAGGC	TTCCAGGTCG	GTGATGTCCG	CGTCGTCTGT	GTTGGTGACG	TGCGGGATGA	60
CCACCCAGTC	GCGGTGCAGG	TTTTTCGATG	GCATAATATC	TGCGTTGCGA	CGTGTAACAC	120
ACTATTGGAG	ACATATCATG	CAAACGCTCA	GCATCCAGCA	CGGTACCCTC	GTCACGATGG	180
ATCAATACCG	CAGAGTCCTT	GGGGATAGCT	GGGTTTACGT	GCAGGATGGA	CGGATCGTCG	240
CGCTCGGAGT	GCACGCCAAG	TCGGTGCCTC	CGCCAGCGGA	TCGGGTGATC	GATGCACGCG	300
GCAAGGTCGT	GTTACCCGGT	TTCATCAATG	CCCACACCCA	TGTGAACCAG	ATCCTCCTGC	360
GCGGAGGGCC	CTCGCACGGG	CGTCAATTCT	ATGACTGGCT	GTTCAACGTT	GTGTATCCGG	420
GACAAAAGGC	GATGAGACCG	GAGGACGTAG	CGGTGGCGGT	GAGGTTGTAT	TGTGCGGAAG	480
CTGTGCGCAG	CGGGATTACG	ACGATCAACG	AAAACGCCGA	TTCGGCCATC	TACCCAGGCA	540
ACATCGAGGC	CGCGATGGCG	GTCTATGGTG	AGGTGGGTGT	GAGGGTCGTC	TACGCCCCGA	600
TGTTCTTTGA	TCGGATGGAC	GGGCGCATT	AAGGGTATGT	GGACGCCTTG	AAGGCTCGCT	660
CTCCCCAAGT	CGAACTGTGC	TCGATCATGG	AGGAAACGGC	TGTGGCCAAA	GATCGGATCA	720
CAGCCCTGTC	AGATCAGTAT	CATGGCACGG	CAGGAGGTCG	TATATCAGTT	TGGCCCGCTC	780
CTGCCACTAC	CACGGCGGTG	ACAGTTGAAG	GAATGCGATG	GGCACAAAGCC	TTCGCCCGTG	840

ATCGGGCGGT AATGTGGACG CTTACATGG CGGAGAGCGA TCATGATGGG CGGATTCATG 900  
 GGATGAGTCC CGCCGAGTAC ATGGAGTGTT ACGGACTCTT GGATGAGCGT CTGCAGGTCC 960  
 CGCATTGCGT GTACTTTGAC CGGAAGGATG TTCGGCTGCT GCACCGCCAC AATGTGAAGG 1020  
 TCGCGTCGCA GGTGTGTAGC AATGCCTACC TCGGCTCAGG GGTGGCCCCC GTGCCAGAGA 1080  
 TGGTGGAGCG CGGCATGGCC GTGGGCATTG GAACAGATAA CGGGAATAGT AATGACTCCG 1140  
 TAAACATGAT CGGAGACATG AAGTTTATGG CCCATATTCA CCGCGCGGTG CATCGGGATG 1200  
 CGGACGTGCT GACCCCAGAG AAGATTCTTG AAATGGCGAC GATCGATGGG GCGCGTTCGT 1260  
 TGGGGATGGA CCACGAGATT GGTTCATCG AAACCGGCAA GCGCGCGGAC CTTATCCTGC 1320  
 TTGACCTGCG TCACCCTCAG ACGACTCCTC ACCATCATTT GGCGGCCACG ATCGTGTTTC 1380  
 AGGCTTACGG CAATGAAGTG GACACTGTCC TGATTGACGG AAACGTTGTG ATGGAGAACC 1440  
 GCTGCTTGAG CTTTCTTCCC CCTGAACGTG AGTTGGCGTT CTTGAGGGA GCGCAGAGCC 1500  
 GCGCCACAGC TATTTTGCAG CGGGCGAACA TGGTGGCTAA CCCAGCTTGG CGCAGCCTCT 1560  
 AGGAAATGAC GCCGTTGCTG CATCCGCCGC CCCTTGAGGA AATCGCTGCC ATCTTGGCGC 1620  
 GGCTCGGATT GGG 1633

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1598 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TCGTGGTTGG TGACGTGCGG GATGACCACC CAGTCGCGGT GCAGGTTTTT CGATGGCATA 60  
 ATATCTGCGT TCGACGTGT AACACACTAT TGGAGACATA TCATGCAAAC GCTCAGCATC 120  
 CAGCACGGTA CCCTCGTCAC GATGGATCAG TACCGCAGAG TCCTTGGGGA TAGCTGGGTT 180  
 CACGTGCAGG ATGGACGGAT CGTCGCGCTC GGAGTGCACG CCGAGTCGGT GCCTCCGCCA 240  
 GCGGATCGGG TGATCGATGC ACGCGGCAAG GTCGTGTTAC CCGGTTTCAT CAATGCCAC 300  
 ACCCATGTGA ACCAGATCCT CCTGCGCGGA GGGCCCTCGC ACGGGCGTCA ATTCTATGAC 360  
 TGGCTGTTCA ACGTTGTGTA TCCGGGACAA AAGCGCATGA GACCGGAGGA CGTAGCGGTG 420

GCGGTGAGGT TGTATTGTGC GGAAGCTGTG CGCAGCGGGA TTACGACGAT CAACGAAAAC 480  
 GCCGATTTCG CCATCTACCC AGGCAACATC GAGGCCGCGA TGGCGGTCTA TGGTGAGGTG 540  
 GGTGTGAGGG TCGTCTACGC CCGCATGTTC TTTGATCGGA TGGACGGGCG CATTCAAGGG 600  
 TATGTGGACG CCTTGAAGGC TCGCTCTCCC CAAGTCGAAC TGTGCTCGAT CATGGAGGAA 660  
 ACGGCTGTGG CCAAAGATCG GATCACAGCC CTGTCTAGATC AGTATCATGG CACGGCAGGA 720  
 GGTCTGATAT CAGTTTGGCC CGCTCCTGCC ACTACCACGG CGGTGACAGT TGAAGGAATG 780  
 CGATGGGCAC AAGCCTTCGC CCGTGATCGG GCGGTAATGT GGACGCTTCA CATGGCGGAG 840  
 AGCGATCATG ATGAGCGGAT TCATGGGATG AGTCCCGCCG AGTACATGGA GTGTCACGGA 900  
 CTCTTGGATG AGCGTCTGCA GGTGCGCAT TGCCTGTACT TTGACCGGAA GGATGTTTCG 960  
 CTGCTGCACC GCCACAATGT GAAGGTCGCG TCGCAGGTTG TGAGCAATGC CTACCTCGGC 1020  
 TCAGGGGTGG CCCCCGTGCC AGAGATGGTG GAGCGCGGCA TGGCCATGGG CATTGGAACA 1080  
 GATAACGGGA ATAGTAATGA CTCCGTAAAC ATGATCGGAG ACATGAAGTT TATGGCCCAT 1140  
 ATTCACCGCG CGGTGCATCG GGATGCGGAC GTGCTGACCC CAGAGAAGAT TCTTGAAATG 1200  
 GCGACGATCG ATGGGGCGCG TTCGTTGGGA ATGGACCACG AGATTGGTTC CATCGAAACC 1260  
 GGCAAGCGCG CGGACCTTAT CCTGCTTGAC CTGCGTCACC CTCAGACGAC TCCTCACCAT 1320  
 CATTTGGCGG CCACGATCGT GTTTCAGGCT TACGGCAATG AGGTGGACAC TGTCCTGATT 1380  
 GACGGAAACG TTGTGATGGA GAACCGCCGC TTGAGCTTTC TTCCCCCTGA ACGTGAGTTG 1440  
 GCGTTCCTTG AGGAAGCGCA GAGCCGCGCC ACAGCTATTT TGCAGCGGGC GAACATGGTG 1500  
 GCTAACCAG CTTGGCGCAG CCTCTAGGAA ATGACGCCGT TGCTGCATCC GCCGCCCTT 1560  
 GAGGAAATCG CTGCCATCTT GGCGCGGCTC GGATTGGG 1598

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1586 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

ACGTGCGGGA TGACCACCCA GTTGCGGTGC AGGTTTTTCG ATGGCGTAAT ATCTGCGTTG 60

CGACGTGTAA CACACTATTG GAGACATATC ATGCAAACGC TCAGCATCCA GCACGGTACC 120  
 CTCGTACAGT TGGATCAGTA CCGCAGAGTC CTTGGGGATA GCTGGGTTCA CGTGCAGGAT 180  
 GGACGGATCG TCGCGCTCGG AGTGCACGCC GAGTCGGTGC CTCCGCCAGC GGATCGGGTG 240  
 ATCGATGCAC GCGGCAAGGT CGTGTTACCC GGTTCATCA ATGCCCACAC CCATGTGAAC 300  
 CAGATCCTCC TGC CGGAGG GCCCTCGCAC GGGCGTCAAT TCTATGACTG GCTGTTC AAC 360  
 GTTGTGTATC CGGGACAAA GCGATGAGA CCTGAGGACG TAGCGGTGGC GGTGAGGTTG 420  
 TATTGTGCGG AAGCTGTGCG CAGCGGGATT ACGACGATCA ACGAAAACGC CGATTCCGCC 480  
 ATCTACCCAG GCAACATCGA GGCCGCGATG GCGGTCTATG GTGAGGTGGG TGTGAGGGTC 540  
 GTCTACGCCC GCATGTTCTT TGATCGGATG GACGGGCGCA TTCAAGGGTA TGTGGACGCC 600  
 TTGAAGGCTC GCTCTCCCA AGTCGAACTG TGCTCGATCA TGGAGGAAAC GGCTGTGGCC 660  
 AAAGATCGGA TCACAGCCCT GTCAGATCAG TATCATGGCA CGGCAGGAGG TCGTATATCA 720  
 GTTTGGCCCG CTCCTGCCAC TACCACGGCG GTGACAGTTG AAGGAATGCG ATGGGCACAA 780  
 GCCTTCGCCC GTGATCGGGC GGTAATGTGG ACGCTTCACA TGGCGGAGAG CGATCATGAT 840  
 GAGCGGATTC ATGGGATGAG TCCCGCCGAG TACATGGAGT GTTACGGACT CTTGGATGAG 900  
 CGTCTGCAGG TCGCGCATTG CGTGTACTTT GACCGGAAGG ATGTTCCGGCT GCTGCACCGC 960  
 CACAATGTGA AGGTCGCGTC GCAGGTTGTG AGCAATGCCT ACCTCGGCTC AGGGGTGGCC 1020  
 CCCGTGCCAG AGATGGTGGG GCGCGGCATG GCCGTGGGCA TTGGAACAGA TAACGGGAAT 1080  
 AGTAATGACT CCGTAAACAT GATCGGAGAC ATGAAGTTTA TGGCCCATAT TCACCGCGCG 1140  
 GTGCATCGGG ATGCGGACGT GCTGACCCCA GAGAAGATTC TTGAAATGGC GACAAATCGAT 1200  
 GGGGCGCGTT CGTTGGGAAT GGACCACGAG ATTGGTTCCA TCGAAACCGG CAAGCGCGCG 1260  
 GACCTTATCC TGCTTGACCT GCGTCACCCT CAGACGACTC CTCACCATCA TTTGGCGGCC 1320  
 ACGATCGTGT TTCAGGCTTA CGGCAATGAG GTGGACACTG TCCTGATTGA CGGAAACGTT 1380  
 GTGATGGAGA ACCGCCGCTT GAGCTTTCTT CCCCTGAAC GTGAGTTGGC GTTCCTTGAG 1440  
 GAAGCGCAGA GCCGCGCCAC AGCTATTTTG CAGCGGGCGA ACATGGTGGC TAACCCAGCT 1500  
 TGGCGCAGCC TCTAGGAAAT GACGCCGTTG CTGCATCCGC TGCCCCTTGA GGAAATCGCT 1560  
 GCCATCTTGG CGCGGCTCGG ATTGGG 1586

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1597 base pairs

(B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CGTGGTTGGT GACGTGGGGG ATGACCACCC AGTCGCGGTG CAGGTTTTTC GATGGCATAA	60
TATCTGCGTT GCGACGTGTA ACACACTATT GGAGACATAT CATGCAAACG CTCAGCATCC	120
AGCACGGTAC CCTCGTCACG ATGGATCAGT ACCGCAGAGT CCTTGGGGAT AGCTGGGTTC	180
ACGTGCAGGA TGGACGGATC GTCGCGCTCG GAGTGACGCG CGAGTCGGTG CCTCCGCCAG	240
CGGATCAGGT GATCGATGCA CGCGGCAAGG TCGTGTTACC CGGTTTCATC AATGCCCACA	300
CCCATGTGAA CCAGATCCTC CTGCGCGGAG GGCCCTCGCA CGGGCGTCAA TTCCATGACT	360
GGCTGTTCAA CGTTGTGTAT CCGGGACAAA AGGCGATGAG ACCGGAGGAC GTAGCGGTGG	420
CGGTGAGGTT GTATTGTGCA GAAGCTGTGC GCAGCGGGAT TACGACGATT AACGAAAACG	480
CCGATTCGGC CATCTACCCA GGCAACATCG AGGCCGCGAT GGCGGTCTAT GGTGAGGTGG	540
GTGTGAGGGT CGTCTACGCC CGCATGTTCT TTGATCGGAT GGACGGGCGC ATTCAAGGCT	600
ATGTGGACGC CTTGAAGGCT CGCTCTCCCC AAGTCGAACT GTGCTCGATC ATGGAGGAAA	660
CGGCTGTGGC CAAAGATCGG ATCACAGCCC TGTCAGATCA GTATCATGGC ACGGCAGGAG	720
GTCGTATATC AGTTTGGCCC GCTCCTGCCA CTACCACGGC GGTGACAGTT GAAGGAATGC	780
GATGGGCACA AGCCTTCGCC CGTGATCGGG CGGTAATGTG GACGCTTCAC ATGGCGGAGA	840
GCGATCATGA TGGGCGGATT CATGGGATGA GTCCCGCCGA GTACATGGAG TGTACGGAC	900
TCTTGATGA GCGTCTGCAG GTCGCGCATT GCGTGACTT TGACCGGAAG GATGTTCGGC	960
TGCTGCACCG CCACAATGTG AAGGTCGCGT CGCAGGTTGT GAGCAATGCC TACCTCGGCT	1020
CAGGGGTGGC CCCCCTGCCA GAGATGGTGG AGCGCGGCAT GGCCGTGGGC ATTGGAACAG	1080
ATAACGGGAA TAGTAATGAC TCCGTAAACA TGATCGGAGA CATGAAGTTT ATGGCCCATA	1140
TTCACCGCGC GGTGCATCGG GATGCGGACG TGCTGACCCC AGAGAAGATT CTTGAAATGG	1200
CAACGATCGA TGGGGCGCGT TCGTTGGGAA TGGACCACGA GATTGGTTCC ATCGAAACCG	1260
GCAAGCGCGC GGACCTTATC CTGCTTGACC TGCCTCACCC TCAGACGACT CCTCACCATC	1320
ATTTGGCGGC CACGATCGTG TTTCAGGCTT ACGGCAATGA GGTGGACACT GTCCTGATTG	1380

10955307.05504

ACGGAAACGT TGTGATGGAG AACCGCCGCT TGAGCTTTCT TCCCCCTGAA CGTGAGTTGG	1440
CGTTCCTTGA GGAAGCGCAG AGCCGCGCCA CAGCTATTTT GCAGCGGGCG AACATGGTGG	1500
CTAACCCAGC TTGGCGCAGC CTCTAGGAAA TGACGCCGTT GCTGCATCCG CCGCCCCTTG	1560
AGGAAATCGC TGCCATCTTG GCGCGGCTCG GATTGGG	1597

## (2) INFORMATION FOR SEQ ID NO:21:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1674 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GTGACCTTGA CGCCGCTCTT TTCGTTCTCT TTGTTGAACT GCACGCGAAT GGCTTCCAGT	60
TGCATGATGT CCGCGTCGTC GTGGTTGGTG ACGTGCGGGA TGACCACCCA GTCGCGGTGC	120
AGGTTTTTCG ATGGCATAAT ATCTGCGTTG CGACGTGTAA CACACTATTG GAGACATATC	180
ATGCAAACGC TCAGCATCCA GCACGGTACC CTCGTCACGA TGGATCAGTA CCGCAGAGTC	240
CTTGGGGATA GCTGGGTTCA CGTGCAGGAT GGACGGATCG TCGCGCTCGG AGTGCACGCC	300
GAGTCGGTGC CTCCGCCAGC GGATCGGGTG ATTGATGCAC GCGGCAAGGT CGTGTTACCC	360
GGTTTCATCA ATGCCCACAC CCATGTGAAC CAGATCCTCC TGC GCGGAGG CCTCGCACGG	420
GCGTCAATTC TATGACTGGC TGTTCAACGT TGTGTATCCG GGACAAAAGG CGATGAGACC	480
GGAGGACGTA GCGGTGGCGG TGAGGTTGTA TTGTGCGGAA GCTGTGCGCA GCGGGATTAC	540
GACGATCAAC GAAAACGCCG ATTCGGCCAT CTACCCAGGC AACATCGAGG CCGCGATGGC	600
GGTCTATGGT GAGGTGGGTG TGAGGGTCGT CTACGCCCCG ATGTTCTTTG ATCGGATGGA	660
CAGGCGCATT CAAGGGTATG TGGACGCCTT GAAGGCTCGC TCTCCCCAAG TCGAACTGTG	720
CTCGATCATG GAGGAAACGG CTGTGGCCAA AGATCGGATC ACAGCCCTGT CAGATCAGTA	780
TCATGGCAGC GCAGGAGGTC GTATATCAGT TTGGCCCCGT CCTGCCACTA CCACGGCGGT	840
GACAGTTGAA GGAATGCGAT GGGCACAAGC CTTGCCCCGT GATCGGGCGG TAATGTGGAC	900
GCTTCACATG GCGGAGAGCG ATCATGATGA GCGGATTCAT GGGATGAGTC CCGCCGAGTA	960
CATGGAGTGT TACGGACTCT TGGATGAGCG TCTGCAGGTC GCGCATTGCG TGTACTTTGA	1020

10550 025500



(2) INFORMATION FOR SEQ ID NO:22:

(A) LENGTH: 496 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met	Gln	Thr	Leu	Ser	Ile	Gln	His	Gly	Thr	Leu	Val	Thr	Met	Asp	Gln
1				5					10					15	
Tyr	Arg	Arg	Val	Leu	Gly	Asp	Ser	Trp	Val	His	Val	Gln	Asp	Gly	Arg
			20					25					30		
Ile	Val	Ala	Leu	Gly	Val	His	Ala	Lys	Ser	Val	Pro	Pro	Pro	Ala	Asp
		35					40					45			
Arg	Val	Ile	Asp	Ala	Arg	Gly	Lys	Val	Val	Leu	Pro	Gly	Phe	Ile	Asn
	50					55					60				
Ala	His	Thr	His	Val	Asn	Gln	Ile	Leu	Leu	Arg	Gly	Gly	Pro	Ser	His
65					70					75					80
Gly	Arg	Gln	Phe	Tyr	Asp	Trp	Leu	Phe	Asn	Val	Val	Tyr	Pro	Gly	Gln
				85					90					95	

Lys Ala Met Arg Pro Glu Asp Val Ala Val Ala Val Arg Leu Tyr Cys  
 100 105 110  
 Ala Glu Ala Val Arg Ser Gly Ile Thr Thr Ile Asn Glu Asn Ala Asp  
 115 120 125  
 Ser Ala Ile Tyr Pro Gly Asn Ile Glu Ala Ala Met Ala Val Tyr Gly  
 130 135 140  
 Glu Val Gly Val Arg Val Val Tyr Ala Arg Met Phe Phe Asp Arg Met  
 145 150 155 160  
 Asp Gly Arg Ile Gln Gly Tyr Val Asp Ala Leu Lys Ala Arg Ser Pro  
 165 170 175  
 Gln Val Glu Leu Cys Ser Ile Met Glu Glu Thr Ala Val Ala Lys Asp  
 180 185 190  
 Arg Ile Thr Ala Leu Ser Asp Gln Tyr His Gly Thr Ala Gly Gly Arg  
 195 200 205  
 Ile Ser Val Trp Pro Ala Pro Ala Thr Thr Thr Ala Val Thr Val Glu  
 210 215 220  
 Gly Met Arg Trp Ala Gln Ala Phe Ala Arg Asp Arg Ala Val Met Trp  
 225 230 235 240  
 Thr Leu His Met Ala Glu Ser Asp His Asp Gly Arg Ile His Gly Met  
 245 250 255  
 Ser Pro Ala Glu Tyr Met Glu Cys Tyr Gly Leu Leu Asp Glu Arg Leu  
 260 265 270  
 Gln Val Ala His Cys Val Tyr Phe Asp Arg Lys Asp Val Arg Leu Leu  
 275 280 285  
 His Arg His Asn Val Lys Val Ala Ser Gln Val Val Ser Asn Ala Tyr  
 290 295 300  
 Leu Gly Ser Gly Val Ala Pro Val Pro Glu Met Val Glu Arg Gly Met  
 305 310 315 320  
 Ala Val Gly Ile Gly Thr Asp Asn Gly Asn Ser Asn Asp Ser Val Asn  
 325 330 335  
 Met Ile Gly Asp Met Lys Phe Met Ala His Ile His Arg Ala Val His  
 340 345 350  
 Arg Asp Ala Asp Val Leu Thr Pro Glu Lys Ile Leu Glu Met Ala Thr  
 355 360 365  
 Ile Asp Gly Ala Arg Ser Leu Gly Met Asp His Glu Ile Gly Ser Ile  
 370 375 380  
 Glu Thr Gly Lys Arg Ala Asp Leu Ile Leu Leu Asp Leu Arg His Pro  
 385 390 395 400

005530 065604  
 105550 105550

Gln Thr Thr Pro His His His Leu Ala Ala Thr Ile Val Phe Gln Ala  
 405 410 415

Tyr Gly Asn Glu Val Asp Thr Val Leu Ile Asp Gly Asn Val Val Met  
 420 425 430

Glu Asn Arg Cys Leu Ser Phe Leu Pro Pro Glu Arg Glu Leu Ala Phe  
 435 440 445

Leu Glu Gly Ala Gln Ser Arg Ala Thr Ala Ile Leu Gln Arg Ala Asn  
 450 455 460

Met Val Ala Asn Pro Ala Trp Arg Ser Leu Glu Met Thr Pro Leu Leu  
 465 470 475 480

His Pro Pro Pro Leu Glu Glu Ile Ala Ala Ile Leu Ala Arg Leu Gly  
 485 490 495

## (2) INFORMATION FOR SEQ ID NO:23:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 496 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Met Gln Thr Leu Ser Ile Gln His Gly Thr Leu Val Thr Met Asp Gln  
 1 5 10 15

Tyr Arg Arg Val Leu Gly Asp Ser Trp Val His Val Gln Asp Gly Arg  
 20 25 30

Ile Val Ala Leu Gly Val His Ala Glu Ser Val Pro Pro Pro Ala Asp  
 35 40 45

Arg Val Ile Asp Ala Arg Gly Lys Val Val Leu Pro Gly Phe Ile Asn  
 50 55 60

Ala His Thr His Val Asn Gln Ile Leu Leu Arg Gly Gly Pro Ser His  
 65 70 75 80

Gly Arg Gln Phe Tyr Asp Trp Leu Phe Asn Val Val Tyr Pro Gly Gln  
 85 90 95

Lys Ala Met Arg Pro Glu Asp Val Ala Val Ala Val Arg Leu Tyr Cys  
 100 105 110

Ala	Glu	Ala	Val	Arg	Ser	Gly	Ile	Thr	Thr	Ile	Asn	Glu	Asn	Ala	Asp
115						120			125						
Ser	Ala	Ile	Tyr	Pro	Gly	Asn	Ile	Glu	Ala	Ala	Met	Ala	Val	Tyr	Gly
130						135			140						
Glu	Val	Gly	Val	Arg	Val	Val	Tyr	Ala	Arg	Met	Phe	Phe	Asp	Arg	Met
145						150			155			160			
Asp	Gly	Arg	Ile	Gln	Gly	Tyr	Val	Asp	Ala	Leu	Lys	Ala	Arg	Ser	Pro
			165						170			175			
Gln	Val	Glu	Leu	Cys	Ser	Ile	Met	Glu	Glu	Thr	Ala	Val	Ala	Lys	Asp
			180						185			190			
Arg	Ile	Thr	Ala	Leu	Ser	Asp	Gln	Tyr	His	Gly	Thr	Ala	Gly	Gly	Arg
			195			200						205			
Ile	Ser	Val	Trp	Pro	Ala	Pro	Ala	Thr	Thr	Thr	Ala	Val	Thr	Val	Glu
210						215						220			
Gly	Met	Arg	Trp	Ala	Gln	Ala	Phe	Ala	Arg	Asp	Arg	Ala	Val	Met	Trp
225						230			235			240			
Thr	Leu	His	Met	Ala	Glu	Ser	Asp	His	Asp	Glu	Arg	Ile	His	Gly	Met
			245						250			255			
Ser	Pro	Ala	Glu	Tyr	Met	Glu	Cys	His	Gly	Leu	Leu	Asp	Glu	Arg	Leu
			260			265						270			
Gln	Val	Ala	His	Cys	Val	Tyr	Phe	Asp	Arg	Lys	Asp	Val	Arg	Leu	Leu
275						280						285			
His	Arg	His	Asn	Val	Lys	Val	Ala	Ser	Gln	Val	Val	Ser	Asn	Ala	Tyr
290						295						300			
Leu	Gly	Ser	Gly	Val	Ala	Pro	Val	Pro	Glu	Met	Val	Glu	Arg	Gly	Met
305						310			315			320			
Ala	Met	Gly	Ile	Gly	Thr	Asp	Asn	Gly	Asn	Ser	Asn	Asp	Ser	Val	Asn
			325						330			335			
Met	Ile	Gly	Asp	Met	Lys	Phe	Met	Ala	His	Ile	His	Arg	Ala	Val	His
			340			345						350			
Arg	Asp	Ala	Asp	Val	Leu	Thr	Pro	Glu	Lys	Ile	Leu	Glu	Met	Ala	Thr
355						360						365			
Ile	Asp	Gly	Ala	Arg	Ser	Leu	Gly	Met	Asp	His	Glu	Ile	Gly	Ser	Ile
370						375						380			
Glu	Thr	Gly	Lys	Arg	Ala	Asp	Leu	Ile	Leu	Leu	Asp	Leu	Arg	His	Pro
385			390						395			400			

[illegible]

(1) SEQUENCE CHARACTERISTICS

- (A) LENGTH: 496 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

SEQUENCE CHARACTERISTICS

- (A) LENGTH: 496 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

It is important to note that the above results are based on the assumption that the data are stationary. If the data are non-stationary, the results may be biased. Therefore, it is important to test for stationarity before conducting the analysis.

Ala	Glu	Ala	Val	Arg	Ser	Gly	Ile	Thr	Thr	Ile	Asn	Glu	Ala	Asp		
115						120						125				
Ser	Ala	Ile	Tyr	Pro	Gly	Asn	Ile	Glu	Ala	Ala	Met	Ala	Val	Tyr	Gly	
130						135						140				
Glu	Val	Gly	Val	Arg	Val	Val	Tyr	Ala	Arg	Met	Phe	Phe	Asp	Arg	Met	
145						150						155			160	
Asp	Gly	Arg	Ile	Gln	Gly	Tyr	Val	Asp	Ala	Leu	Lys	Ala	Arg	Ser	Pro	
165						170						175				
Gln	Val	Glu	Leu	Cys	Ser	Ile	Met	Glu	Glu	Thr	Ala	Val	Ala	Lys	Asp	
180						185						190				
Arg	Ile	Thr	Ala	Leu	Ser	Asp	Gln	Tyr	His	Gly	Thr	Ala	Gly	Gly	Arg	
195						200						205				
Ile	Ser	Val	Trp	Pro	Ala	Pro	Ala	Thr	Thr	Thr	Ala	Val	Thr	Val	Glu	
210						215						220				
Gly	Met	Arg	Trp	Ala	Gln	Ala	Phe	Ala	Arg	Asp	Arg	Ala	Val	Met	Trp	
225						230						235			240	
Thr	Leu	His	Met	Ala	Glu	Ser	Asp	His	Asp	Glu	Arg	Ile	His	Gly	Met	
245						250						255				
Ser	Pro	Ala	Glu	Tyr	Met	Glu	Cys	Tyr	Gly	Leu	Leu	Asp	Glu	Arg	Leu	
260						265						270				
Gln	Val	Ala	His	Cys	Val	Tyr	Phe	Asp	Arg	Lys	Asp	Val	Arg	Leu	Leu	
275						280						285				
His	Arg	His	Asn	Val	Lys	Val	Ala	Ser	Gln	Val	Val	Ser	Asn	Ala	Tyr	
290						295						300				
Leu	Gly	Ser	Gly	Val	Ala	Pro	Val	Pro	Glu	Met	Val	Glu	Arg	Gly	Met	
305						310						315			320	
Ala	Val	Gly	Ile	Gly	Thr	Asp	Asn	Gly	Asn	Ser	Asn	Asp	Ser	Val	Asn	
325						330						335				
Met	Ile	Gly	Asp	Met	Lys	Phe	Met	Ala	His	Ile	His	Arg	Ala	Val	His	
340						345						350				
Arg	Asp	Ala	Asp	Val	Leu	Thr	Pro	Glu	Lys	Ile	Leu	Glu	Met	Ala	Thr	
355						360						365				
Ile	Asp	Gly	Ala	Arg	Ser	Leu	Gly	Met	Asp	His	Glu	Ile	Gly	Ser	Ile	
370						375						380				
Glu	Thr	Gly	Lys	Arg	Ala	Asp	Leu	Ile	Leu	Leu	Asp	Leu	Arg	His	Pro	
385						390						395			400	
Gln	Thr	Thr	Pro	His	His	His	Leu	Ala	Ala	Thr	Ile	Val	Phe	Gln	Ala	
405						410						415				

Tyr Gly Asn Glu Val Asp Thr Val Leu Ile Asp Gly Asn Val Val Met  
 420 425 430  
 Glu Asn Arg Arg Leu Ser Phe Leu Pro Pro Glu Arg Glu Leu Ala Phe  
 435 440 445  
 Leu Glu Glu Ala Gln Ser Arg Ala Thr Ala Ile Leu Gln Arg Ala Asn  
 450 455 460  
 Met Val Ala Asn Pro Ala Trp Arg Ser Leu Glu Met Thr Pro Leu Leu  
 465 470 475 480  
 His Pro Leu Pro Leu Glu Glu Ile Ala Ala Ile Leu Ala Arg Leu Gly  
 485 490 495

## (2) INFORMATION FOR SEQ ID NO:25:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 496 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Met Gln Thr Leu Ser Ile Gln His Gly Thr Leu Val Thr Met Asp Gln  
 1 5 10 15  
 Tyr Arg Arg Val Leu Gly Asp Ser Trp Val His Val Gln Asp Gly Arg  
 20 25 30  
 Ile Val Ala Leu Gly Val His Ala Glu Ser Val Pro Pro Pro Ala Asp  
 35 40 45  
 Gln Val Ile Asp Ala Arg Gly Lys Val Val Leu Pro Gly Phe Ile Asn  
 50 55 60  
 Ala His Thr His Val Asn Gln Ile Leu Leu Arg Gly Gly Pro Ser His  
 65 70 75 80  
 Gly Arg Gln Phe His Asp Trp Leu Phe Asn Val Val Tyr Pro Gly Gln  
 85 90 95  
 Lys Ala Met Arg Pro Glu Asp Val Ala Val Ala Val Arg Leu Tyr Cys  
 100 105 110  
 Ala Glu Ala Val Arg Ser Gly Ile Thr Thr Ile Asn Glu Asn Ala Asp  
 115 120 125

Ser	Ala	Ile	Tyr	Pro	Gly	Asn	Ile	Glu	Ala	Ala	Met	Ala	Val	Tyr	Gly	
	130					135					140					
Glu	Val	Gly	Val	Arg	Val	Val	Tyr	Ala	Arg	Met	Phe	Phe	Asp	Arg	Met	
145					150					155					160	
Asp	Gly	Arg	Ile	Gln	Gly	Tyr	Val	Asp	Ala	Leu	Lys	Ala	Arg	Ser	Pro	
				165					170					175		
Gln	Val	Glu	Leu	Cys	Ser	Ile	Met	Glu	Glu	Thr	Ala	Val	Ala	Lys	Asp	
			180					185					190			
Arg	Ile	Thr	Ala	Leu	Ser	Asp	Gln	Tyr	His	Gly	Thr	Ala	Gly	Gly	Arg	
		195					200					205				
Ile	Ser	Val	Trp	Pro	Ala	Pro	Ala	Thr	Thr	Thr	Ala	Val	Thr	Val	Glu	
	210					215					220					
Gly	Met	Arg	Trp	Ala	Gln	Ala	Phe	Ala	Arg	Asp	Arg	Ala	Val	Met	Trp	
225					230					235					240	
Thr	Leu	His	Met	Ala	Glu	Ser	Asp	His	Asp	Gly	Arg	Ile	His	Gly	Met	
				245					250					255		
Ser	Pro	Ala	Glu	Tyr	Met	Glu	Cys	Tyr	Gly	Leu	Leu	Asp	Glu	Arg	Leu	
			260					265					270			
Gln	Val	Ala	His	Cys	Val	Tyr	Phe	Asp	Arg	Lys	Asp	Val	Arg	Leu	Leu	
		275					280					285				
His	Arg	His	Asn	Val	Lys	Val	Ala	Ser	Gln	Val	Val	Ser	Asn	Ala	Tyr	
	290					295					300					
Leu	Gly	Ser	Gly	Val	Ala	Pro	Val	Pro	Glu	Met	Val	Glu	Arg	Gly	Met	
305					310					315					320	
Ala	Val	Gly	Ile	Gly	Thr	Asp	Asn	Gly	Asn	Ser	Asn	Asp	Ser	Val	Asn	
				325					330					335		
Met	Ile	Gly	Asp	Met	Lys	Phe	Met	Ala	His	Ile	His	Arg	Ala	Val	His	
			340					345					350			
Arg	Asp	Ala	Asp	Val	Leu	Thr	Pro	Glu	Lys	Ile	Leu	Glu	Met	Ala	Thr	
		355					360					365				
Ile	Asp	Gly	Ala	Arg	Ser	Leu	Gly	Met	Asp	His	Glu	Ile	Gly	Ser	Ile	
	370					375					380					
Glu	Thr	Gly	Lys	Arg	Ala	Asp	Leu	Ile	Leu	Leu	Asp	Leu	Arg	His	Pro	
385					390					395					400	
Gln	Thr	Thr	Pro	His	His	His	Leu	Ala	Ala	Thr	Ile	Val	Phe	Gln	Ala	
				405					410					415		
Tyr	Gly	Asn	Glu	Val	Asp	Thr	Val	Leu	Ile	Asp	Gly	Asn	Val	Val	Met	
			420					425					430			



Glu Asn Arg Arg Leu Ser Phe Leu Pro Pro Glu Arg Glu Leu Ala Phe  
                   435                                  440                                  445  
  
 Leu Glu Glu Ala Gln Ser Arg Ala Thr Ala Ile Leu Gln Arg Ala Asn  
                   450                                  455                                  460  
  
 Met Val Ala Asn Pro Ala Trp Arg Ser Leu Glu Met Thr Pro Leu Leu  
                   465                                  470                                  475                                  480  
  
 His Pro Pro Pro Leu Glu Glu Ile Ala Ala Ile Leu Ala Arg Leu Gly  
                                   485                                  490                                  495

## (2) INFORMATION FOR SEQ ID NO:26:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 496 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Met Gln Thr Leu Ser Ile Gln His Gly Thr Leu Val Thr Met Asp Gln  
 1                                  5                                  10                                  15  
  
 Tyr Arg Arg Val Leu Gly Asp Ser Trp Val His Val Gln Asp Gly Arg  
                   20                                  25                                  30  
  
 Ile Val Ala Leu Gly Val His Ala Glu Ser Val Pro Pro Pro Ala Asp  
                   35                                  40                                  45  
  
 Arg Val Ile Asp Ala Arg Gly Lys Val Val Leu Pro Gly Phe Ile Asn  
                   50                                  55                                  60  
  
 Ala His Thr His Val Asn Gln Ile Leu Leu Arg Gly Gly Pro Ser His  
                   65                                  70                                  75                                  80  
  
 Gly Arg Gln Phe Tyr Asp Trp Leu Phe Asn Val Val Tyr Pro Gly Gln  
                   85                                  90                                  95  
  
 Lys Ala Met Arg Pro Glu Asp Val Ala Val Ala Val Arg Leu Tyr Cys  
                   100                                  105                                  110  
  
 Ala Glu Ala Val Arg Ser Gly Ile Thr Thr Ile Asn Glu Asn Ala Asp  
                   115                                  120                                  125  
  
 Ser Ala Ile Tyr Pro Gly Asn Ile Glu Ala Ala Met Ala Val Tyr Gly  
                   130                                  135                                  140

Glu	Val	Gly	Val	Arg	Val	Val	Tyr	Ala	Arg	Met	Phe	Phe	Asp	Arg	Met
145					150					155					160
Asp	Arg	Arg	Ile	Gln	Gly	Tyr	Val	Asp	Ala	Leu	Lys	Ala	Arg	Ser	Pro
			165						170					175	
Gln	Val	Glu	Leu	Cys	Ser	Ile	Met	Glu	Glu	Thr	Ala	Val	Ala	Lys	Asp
			180					185					190		
Arg	Ile	Thr	Ala	Leu	Ser	Asp	Gln	Tyr	His	Gly	Thr	Ala	Gly	Gly	Arg
		195					200					205			
Ile	Ser	Val	Trp	Pro	Ala	Pro	Ala	Thr	Thr	Thr	Ala	Val	Thr	Val	Glu
	210					215					220				
Gly	Met	Arg	Trp	Ala	Gln	Ala	Phe	Ala	Arg	Asp	Arg	Ala	Val	Met	Trp
225					230					235					240
Thr	Leu	His	Met	Ala	Glu	Ser	Asp	His	Asp	Glu	Arg	Ile	His	Gly	Met
				245					250					255	
Ser	Pro	Ala	Glu	Tyr	Met	Glu	Cys	Tyr	Gly	Leu	Leu	Asp	Glu	Arg	Leu
			260					265					270		
Gln	Val	Ala	His	Cys	Val	Tyr	Phe	Asp	Arg	Lys	Asp	Ile	Arg	Leu	Leu
		275					280					285			
His	Arg	His	Asn	Val	Lys	Val	Ala	Ser	Gln	Ala	Val	Ser	Asn	Ala	Tyr
	290					295					300				
Leu	Gly	Ser	Gly	Val	Ala	Pro	Val	Pro	Glu	Met	Val	Glu	Arg	Gly	Met
305					310					315					320
Ala	Val	Gly	Ile	Gly	Thr	Asp	Asn	Gly	Asn	Ser	Asn	Asp	Ser	Val	Asn
				325					330					335	
Met	Ile	Gly	Asp	Met	Lys	Phe	Met	Ala	His	Ile	His	Arg	Ala	Val	His
			340					345					350		
Arg	Asp	Ala	Asp	Val	Leu	Thr	Pro	Glu	Lys	Ile	Leu	Glu	Met	Ala	Thr
		355					360					365			
Ile	Asp	Gly	Ala	Arg	Ser	Leu	Gly	Met	Asp	His	Glu	Ile	Gly	Ser	Ile
	370					375					380				
Glu	Thr	Gly	Lys	Arg	Ala	Asp	Leu	Ile	Leu	Leu	Asp	Leu	Arg	His	Pro
385					390					395					400
Gln	Thr	Thr	Pro	His	His	His	Leu	Ala	Ala	Thr	Ile	Val	Phe	Gln	Ala
				405					410					415	
Tyr	Gly	Asn	Glu	Val	Asp	Thr	Val	Leu	Ile	Asp	Gly	Asn	Val	Val	Met
			420					425					430		
Glu	Asn	Arg	Arg	Leu	Ser	Phe	Leu	Pro	Pro	Glu	Arg	Glu	Leu	Ala	Phe
		435					440					445			

Met Val Ala Asn Pro Ala Trp Arg Ser Leu Glu Met Thr Pro Leu Leu  
465 470 475 480

His Pro Pro Pro Leu Glu Glu Ile Ala Ala Ile Leu Ala Gln Leu Gly  
485 490 495

[illegible]